BEST AVAILABLE COPY

Figure 1A

FIG 1 B

```
   1 CTCGAGAAAT CATAAAAAAT TTATTTGCTT TGTGAGCGGA TAACAATTAT AATAGATTCA
  61 ATTGTGAGCG GATAACAATT TCACACAGAA TTCATTAAAG AGGAGAAATT AACTATGGCA
 121 CTTAGTGGGA TCCGCATGCG AGCTCGGTAC CCCGGGGGTG GCAGCGGTTC TGGCGCAGCA
 181 GCGGAAATCA GTGGTCACAT CGTACGTTCC CCGATGGTTG GTACTTTCTA CCGCACCCCA
 241 AGCCCGGACG CAAAAGCGTT CATCGAAGTG GGTCAGAAAG TCAACGTGGG CGATACCCTG
 301 TGCATCGTTG AAGCCATGAA AATGATGAAC CAGATCGAAG CGGACAAATC CGGTACCGTG
 361 AAAGCAATTC TGGTCGAAAG TGGACAACCG GTAGAATTTG ACGAGCCGCT GGTCGTCATC
 421 GAGGGTGGCA GCGGTTCTGG CCACCATCAC CATCACCATA AGCTTAATTA GCTGAGCTTG
 481 GACTCCTGTT GATAGATCCA GTAATGACCT CAGAACTCCA TCTGGATTTG TTCAGAACGC
 541 TCGGTTGCCG CCGGGCGTTT TTTATTGGTG AGAATCCAAG CTAGCTTGGC GAGATTTTCA
 601 GGAGCTAAGG AAGCTAAAAT GGAGAAAAAA ATCACTGGAT ATACCACCGT TGATATATCC
 661 CAATGGCATC GTAAAGAACA TTTTGAGGCA TTTCAGTCAG TTGCTCAATG TACCTATAAC
 721 CAGACCGTTC AGCTGGATAT TACGGCCTTT TTAAAGACCG TAAAGAAAAA TAAGCACAAG
 781 TTTTATCCGG CCTTTATTCA CATTCTTGCC CGCCTGATGA ATGCTCATCC GGAATTTCGT
 841 ATGGCAATGA AGACGGTGA GCTGGTGATA TGGGATAGTG TTCACCCTTG TTACACCGTT
 901 TTCCATGAGC AAACTGAAAC GTTTTCATCG CTCTGGAGTG AATACCACGA CGATTTCCGG
 961 CAGTTTCTAC ACATATATTC GCAAGATGTG GCGTGTTACG GTGAAAACCT GGCCTATTTC
1021 CCTAAAGGGT TTATTGAGAA TATGTTTTTC GTCTCAGCCA ATCCCTGGGT GAGTTTCACC
1081 AGTTTTGATT TAAACGTGGC CAATATGGAC AACTTCTTCG CCCCCGTTTT CACCATGGGC
1141 AAATATTATA CGCAAGGCGA CAAGGTGCTG ATGCCGCTGG CGATTCAGGT TCATCATGCC
1201 GTTTGTGATG GCTTCCATGT CGGCAGAATG CTTAATGAAT TACAACAGTA CTGCGATGAG
1261 TGGCAGGGCG GGGCGTAATT TTTTTAAGGC AGTTATTGGT GCCCTTAAAC GCCTGGGGTA
1321 ATGACTCTCT AGCTTGAGGC ATCAAATAAA ACGAAAGGCT CAGTCGAAAG ACTGGGCCTT
1381 TCGTTTTATC TGTTGTTTGT CGGTGAACGC TCTCCTGAGT AGGACAAATC CGCCCTCTAG
1441 ATTACGTGCA GTCGATGATA AGCTGTCAAA CATGAGAATT GTGCCTAATG AGTGAGCTAA
1501 CTTACATTAA TTGCGTTGCG CTCACTGCCC GCTTTCCAGT CGGGAAACCT GTCGTGCCAG
1561 CTGCATTAAT GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TGCGTATTGG GCGCCAGGGT
1621 GGTTTTTCTT TTCACCAGTG AGACGGGCAA CAGCTGATTG CCCTTCACCG CCTGGCCCTG
1681 AGAGAGTTGC AGCAAGCGGT CCACGCTGGT TTGCCCCAGC AGGCGAAAAT CCTGTTTGAT
1741 GGTGGTTAAC GGCGGGATAT AACATGAGCT GTCTTCGGTA TCGTCGTATC CCACTACCGA
1801 GATATCCGCA CCAACGCGCA GCCCGGACTC GGTAATGGCG CGCATTGCGC CCAGCGCCAT
1861 CTGATCGTTG GCAACCAGCA TCGCAGTGGG AACGATGCCC TCATTCAGCA TTTGCATGGT
1921 TTGTTGAAAA CCGGACATGG CACTCCAGTC GCCTTCCCGT TCCGCTATCG GCTGAATTTG
1981 ATTGCGAGTG AGATATTTAT GCCAGCCAGC CAGACGCAGA CGCGCCGAGA CAGAACTTAA
2041 TGGGCCCGCT AACAGCGCGA TTTGCTGGTG ACCCAATGCG ACCAGATGCT CCACGCCCAG
2101 TCGCGTACCG TCTTCATGGG AGAAAATAAT ACTGTTGATG GGTGTCTGGT CAGAGACATC
2161 AAGAAATAAC GCCGGAACAT TAGTGCAGGC AGCTTCCACA GCAATGGCAT CCTGGTCATC
2221 CAGCGGATAG TTAATGATCA GCCCACTGAC GCGTTGCGCG AGAAGATTGT GCACCGCCGC
2281 TTTACAGGCT TCGACGCCGC TTCGTTCTAC CATCGACACC ACCACGCTGG CACCCAGTTG
2341 ATCGGCGCGA GATTTAATCG CCGCGACAAT TTGCGACGGC GCGTGCAGGG CCAGACTGGA
2401 GGTGGCAACG CCAATCAGCA ACGACTGTTT GCCCGCCAGT TGTTGTGCCA CGCGGTTGGG
2461 AATGTAATTC AGCTCCGCCA TCGCCGCTTC CACTTTTTCC CGCGTTTTCG CAGAAACGTG
2521 GCTGGCCTGG TTCACCACGC GGGAAACGGT CTGATAAGAG ACACCGGCAT ACTCTGCGAC
2581 ATCGTATAAC GTTACTGGTT TCACATTCAC CACCCTGAAT TGACTCTCTT CCGGGCGCTA
2641 TCATGCCATA CCGCGAAAGG TTTTGCACCA TTCGATGGTG TCGGAATTTC GGGCAGCGTT
2701 GGGTCCTGGC CACGGGTGCG CATGATCTAG AGCTGCCTCG CGCGTTTCGG TGATGACGGT
2761 GAAAACCTCT GACACATGCA GCTCCCGGAG ACGGTCACAG CTTGTCTGTA AGCGGATGCC
2821 GGGAGCAGAC AAGCCCGTCA GGGCGCGTCA GCGGGTGTTG GCGGGTGTCG GGGCGCAGCC
2881 ATGACCCAGT CACGTAGCGA TAGCGGAGTG TATACTGGCT TAACTATGCG GCATCAGAGC
2941 AGATTGTACT GAGAGTGCAC CATATGCGGT GTGAAATACC GCACAGATGC GTAAGGAGAA
3001 AATACCGCAT CAGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC
3061 GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG
3121 GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA
3181 AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC
3241 GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC
3301 CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG
3361 CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT
3421 CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC
3481 GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
3541 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
3601 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
3661 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA
3721 CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG
3781 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
3841 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA
3901 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
```

```
3961 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG
4021 TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
4081 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
4141 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
4201 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
4261 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
4321 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
4381 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
4441 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
4501 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
4561 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
4621 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
4681 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
4741 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
4801 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
4861 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC
4921 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
4981 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT TCAC
```

Figure 1B

```
        Dra III              Sph I                 Sma I
115 ATGGCA CTTAGTGGGA TCCGCATGCG AGCTCGGTAC CCCGGGGGTG GCAGC
        TACCGT GAATCACCCT AGGCGTACGC TCGAGCCATG GGGCCCCCAC CGTCG
```
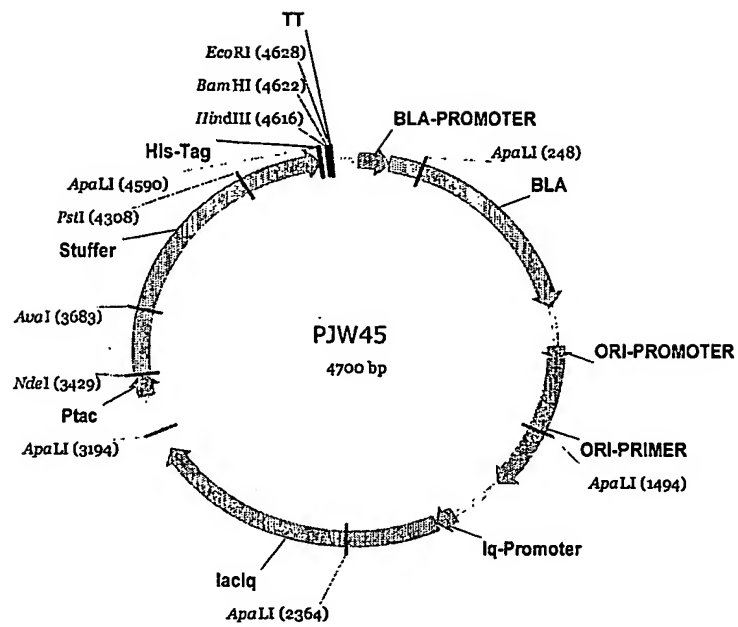
Figure 1C

Figure 2A

```
   1 CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC
  61 ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA
 121 AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT
 181 TTTGCCTTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
 241 AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA
 301 GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG
 361 CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
 421 AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG
 481 TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
 541 TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG
 601 TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG
 661 ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC
 721 TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC
 781 CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG
 841 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG
 901 TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG
 961 AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC
1021 TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG
1081 ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
1141 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC
1201 AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
1261 TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT
1321 AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC.
1381 TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT
1441 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC
1501 AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCATTGAG
1561 AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG
1621 GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG
1681 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
1741 GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT
1801 TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT
1861 TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG
1921 AGGAAGCCCA GGACCCAACG CTGCCCGAAA TTCCGACACC ATCGAATGGT GCAAAACCTT
1981 TCGCGGTATG GCATGATAGC GCCCGGAAGA GAGTCAATTC AGGGTGGTGA ATGTGAAACC
2041 AGTAACGTTA TACGATGTCG CAGAGTATGC CGGTGTCTCT TATCAGACCG TTTCCCGCGT
2101 GGTGAACCAG GCCAGCCACG TTTCTGCGAA AACGCGGGAA AAAGTGGAAG CGGCGATGGC
2161 GGAGCTGAAT TACATTCCCA ACCGCGTGGC ACAACAACTG GCGGGCAAAC AGTCGTTGCT
2221 GATTGGCGTT GCCACCTCCA GTCTGGCCCT GCACGCGCCG TCGCAAATTG TCGCGGCGAT
2281 TAAATCTCGC GCCGATCAAC TGGGTGCCAG CGTGGTGGTG TCGATGGTAG AACGAAGCGG
2341 CGTCGAAGCC TGTAAAGCGG CGGTGCACAA TCTTCTCGCG CAACGCGTCA GTGGGCTGAT
2401 CATTAACTAT CCGCTGGATG ACCAGGATGC CATTGCTGTG GAAGCTGCCT GCACTAATGT
2461 TCCGGCGTTA TTTCTTGATG TCTCTGACCA GACACCCATC AACAGTATTA TTTTCTCCCA
2521 TGAAGACGGT ACGCGACTGG GCGTGGAGCA TCTGGTCGCA TTGGGTCACC AGCAAATCGC
```

```
2581 GCTGTTAGCG GGCCCATTAA GTTCTGTCTC GGCGCGTCTG CGTCTGGCTG GCTGGCATAA
2641 ATATCTCACT CGCAATCAAA TTCAGCCGAT AGCGGAACGG GAAGGCGACT GGAGTGCCAT
2701 GTCCGGTTTT CAACAAACCA TGCAAATGCT GAATGAGGGC ATCGTTCCCA CTGCGATGCT
2761 GGTTGCCAAC GATCAGATGG CGCTGGGCGC AATGCGCGCC ATTACCGAGT CCGGGCTGCG
2821 CGTTGGTGCG GATATCTCGG TAGTGGGATA CGACGATACC GAAGACAGCT CATGTTATAT
2881 CCCGCCGTTA ACCACCATCA AACAGGATTT TCGCCTGCTG GGGCAAACCA GCGTGGACCG
2941 CTTGCTGCAA CTCTCTCAGG GCCAGGCGGT GAAGGGCAAT CAGCTGTTGC CCGTCTCACT
3001 GGTGAAAAGA AAAACCACCC TGGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC
3061 CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA
3121 ACGCAATTAA TGTGAGTTAG CTCACTCATT AGGCACAATT CTCATGTTTG ACAGCTTATC
3181 ATCGACTGCA CGGTGCACCA ATGCTTCTGG CGTCAGGCAG CCATCGGAAG CTGTGGTATG
3241 GCTGTGCAGG TCGTAAATCA CTGCATAATT CGTGTCGCTC AAGGCGCACT CCCGTTCTGG
3301 ATAATGTTTT TTGCGCCGAC ATCATAACGG TTCTGGCAAA TATTCTGAAA TGAGCTGTTG
3361 ACAATTAATC ATCGGCTCGT ATAATGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG
3421 GAAACACATA TGAACGACTT TCATCGCGAT ACGTGGGCGG AAGTGGATTT GGACGCCATT
3481 TACGACAATG TGGCGAATTT GCGCCGTTTG CTGCCGGACG ACACGCACAT TATGGCGGTC
3541 GTGAAGGCGA ACGCCTATGG ACATGGGGAT GTGCAGGTGG CAAGGACAGC GCTCGAAGCG
3601 GGGGCCTCCC GCCTGGCGGT TGCCTTTTTG GATGAGGCGC TCGCTTTAAG GGAAAAAGGA
3661 ATCGAAGCGC CGATTCTAGT TCTCGGGGCT TCCCGTCCAG CTGATGCGGC GCTGGCCGCC
3721 CAGCAGCGCA TTGCCCTGAC CGTGTTCCGC TCCGACTGGT TGGAAGAAGC GTCCGCCCTT
3781 TACAGCGGCC CTATTCCTAT TCATTTCCAT TTGAAAATGG ACACCGGCAT GGGACGGCTT
3841 GGAGTGAAAG ACGAGGAGGA GACGAAACGA ATCGCAGCGC TGATTGAGCG CCATCCGCAT
3901 TTTGTGCTTG AAGGGGCGTA CACGCATTTT GCGACTGCGG ATGAGGTGAA CACCGATTAT
3961 TTTTCCTATC AGTATACCCG TTTTTTGCAC ATGCTCGAAT GGCTGCCGTC GCGCCCGCCG
4021 CTCGTCCATT GCGCCAACAG CGCAGCGTCG CTCCGTTTCC CTGACCGGAC GTTCAATATG
4081 GTCCGCTTCG GCATTGCCAT GTATGGGCTT GCCCCGTCGC CCGGCATCAA GCCGCTGCTG
4141 CCGTATCCAT TAAAGAAGC ATTTTCGCTC CATAGCCGCC TCGTACACGT CAAAAAACTG
4201 CAACCAGGCG AAAAGGTGAG CTATGGTGCG ACGTACACTG CGCAGACGGA GGAGTGGATC
4261 GGGACGATTC CGATCGGCTA TGCGGACGGC TGGCTCCGCC GCCTGCAGCA CTTTCATGTC
4321 CTTGTTGACG GACAAAAGGC GCCGATTGTC GGCCGCATTT GCATGGACCA GTGCATGATC
4381 CGCCTGCCTG GGCCGCTGCC GGTCGGCACG AAGGTGACAC TGATTGGTCG CCAGGGGGAC
4441 GAGGTAATTT CCATTGATGA TGTCGCTCGC CATTTGGAAA CGATCAACTA CGAAGTGCCT
4501 TGCACGATCA GCTATCGAGT GCCCCGTATT TTTTTCCGCC ATAAGCGTAT AATGGAAGTG
4561 AGAAACGCCA TTGGCCGCGG GGAAAGCAGT GCACATCACC ATCACCATCA CTAAAAGCTT
4621 GGATCCGAAT TCAGCCCGCC TAATGAGCGG GCTTTTTTTT GAACAAAATT AGCTTGGCTG
4681 TTTTGGCGGA TGAGAGAAGA
```

Figure 2B

```
1      ATGGCTCTCA TCCCAGACTT GGCCATGGAA ACCTGGCTTC TCCTGGCTGT
CAGCCTGGTG
61     CTCCTCTATC TATATGGAAC CCATTCACAT GGACTTTTTA AGAAGCTTGG
AATTCCAGGG
121    CCCACACCTC TGCCTTTTTT GGGAAATATT TTGTCCTACC ATAAGGGCTT
TTGTATGTTT
181    GACATGGAAT GTCATAAAAA GTATGGAAAA GTGTGGGGCT TTTATGATGG
TCAACAGCCT
241    GTGCTGGCTA TCACAGATCC TGACATGATC AAAACAGTGC TAGTGAAAGA
ATGTTATTCT
301    GTCTTCACAA ACCGGAGGCC TTTTGGTCCA GTGGGATTTA TGAAAAGTGC
CATCTCTATA
361    GCTGAGGATG AAGAATGGAA GAGATTACGA TCATTGCTGT CTCCAACCTT
CACCAGTGGA
421    AAACTCAAGG AGATGGTCCC TATCATTGCC CAGTATGGAG ATGTGTTGGT
GAGAAATCTG
481    AGGCGGGAAG CAGAGACAGG CAAGCCTGTC ACCTTGAAAG ACGTCTTTGG
GGCCTACAGC
541    ATGGATGTGA TCACTAGCAC ATCATTTGGA GTGAACATCG ACTCTCTCAA
CAATCCACAA
601    GACCCCTTTG TGGAAAACAC CAAGAAGCTT TTAAGATTTG ATTTTTTGGA
TCCATTCTTT
661    CTCTCAATAA CAGTCTTTCC ATTCCTCATC CCAATTCTTG AAGTATTAAA
TATCTGTGTG
721    TTTCCAAGAG AAGTTACAAA TTTTTTAAGA AAATCTGTAA AAAGGATGAA
AGAAAGTCGC
781    CTCGAAGATA CACAAAAGCA CCGAGTGGAT TTCCTTCAGC TGATGATTGA
CTCTCAGAAT
841    TCAAAAGAAA CTGAGTCCCA CAAAGCTCTG TCCGATCTGG AGCTCGTGGC
CCAATCAATT
901    ATCTTTATTT TTGCTGGCTA TGAAACCACG AGCAGTGTTC TCTCCTTCAT
TATGTATGAA
961    CTGGCCACTC ACCCTGATGT CCAGCAGAAA CTGCAGGAGG AAATTGATGC
AGTTTTACCC
1021   AATAAGGCAC CACCCACCTA TGATACTGTG CTACAGATGG AGTATCTTGA
CATGGTGGTG
1081   AATGAAACGC TCAGATTATT CCCAATTGCT ATGAGACTTG AGAGGGTCTG
CAAAAAAGAT
1141   GTTGAGATCA ATGGGATGTT CATTCCCAAA GGGGTGGTGG TGATGATTCC
AAGCTATGCT
1201   CTTCACCGTG ACCCAAAGTA CTGGACAGAG CCTGAGAAGT TCCTCCCTGA
AAGATTCAGC
1261   AAGAAGAACA AGGACAACAT AGATCCTTAC ATATACACAC CCTTTGGAAG
TGGACCCAGA
1321   AACTGCATTG GCATGAGGTT TGCTCTCATG AACATGAAAC TTGCTCTAAT
CAGAGTCCTT
1381   CAGAACTTCT CCTTCAAACC TTGTAAAGAA ACACAGATCC CCCTGAAATT
AAGCTTAGGA
1441   GGACTTCTTC AACCAGAAAA ACCCGTTGTT CTAAAGGTTG AGTCAAGGGA
TGGCACCGTA
1501   AGTGGAGCCT GA
```

Figure 3A

```
1    MALIPDLAME TWLLLAVSLV LLYLYGTHSH GLFKKLGIPG PTPLPFLGNI LSYHKGFCMF
61   DMECHKKYGK VWGFYDGQQP VLAITDPDMI KTVLVKECYS VFTNRRPFGP VGFMKSAISI
121  AEDEEWKRLR SLLSPTFTSG KLKEMVPIIA QYGDVLVRNL RREAETGKPV TLKDVFGAYS
181  MDVITSTSFG VNIDSLNNPQ DPFVENTKKL LRFDFLDPFF LSITVFPFLI PILEVLNICV
241  FPREVTNFLR KSVKRMKESR LEDTQKHRVD FLQLMIDSQN SKETESHKAL SDLELVAQSI
301  IFIFAGYETT SSVLSFIMYE LATHPDVQQK LQEEIDAVLP NKAPPTYDTV LQMEYLDMVV
361  NETLRLFPIA MRLERVCKKD VEINGMFIPK GVVVMIPSYA LHRDPKYWTE PEKFLPERFS
421  KKNKDNIDPY IYTPFGSGPR NCIGMRFALM NMKLALIRVL QNFSFKPCKE TQIPLKLSLG
481  GLLQPEKPVV LKVESRDGTV SGA*
```

Figure 3B

```
   1 ATGGATTCTC TTGTGGTCCT TGTGCTCTGT CTCTCATGTT TGCTTCTCCT TTCACTCTGG
  61 AGACAGAGCT CTGGGAGAGG AAAACTCCCT CCTGGCCCCA CTCCTCTCCC AGTGATTGGA
 121 AATATCCTAC AGATAGGTAT TAAGGACATC AGCAAATCCT TAACCAATCT CTCAAAGGTC
 181 TATGGCCCGG TGTTCACTCT GTATTTTGGC CTGAAACCCA TAGTGGTGCT GCATGGATAT
 241 GAAGCAGTGA AGGAAGCCCT GATTGATCTT GGAGAGGAGT TTTCTGGAAG AGGCATTTTC
 301 CCACTGGCTG AAAGAGCTAA CAGAGGATTT GGAATTGTTT TCAGCAATGG AAAGAAATGG
 361 AAGGAGATCC GGCGTTTCTC CCTCATGACG CTGCGGAATT TTGGGATGGG GAAGAGGAGC
 421 ATTGAGGACC GTGTTCAAGA GGAAGCCCGC TGCCTTGTGG AGGAGTTGAG AAAAACCAAG
 481 GCCTCACCCT GTGATCCCAC TTTCATCCTG GGCTGTGCTC CCTGCAATGT GATCTGCTCC
 541 ATTATTTTCC ATAAACGTTT TGATTATAAA GATCAGCAAT TTCTTAACTT AATGGAAAAG
 601 TTGAATGAAA ACATCAAGAT TTTGAGCAGC CCCTGGATCC AGATCTGCAA TAATTTTTCT
 661 CCTATCATTG ATTACTTCCC GGGAACTCAC AACAAATTAC TTAAAAACGT TGCTTTTATG
 721 AAAAGTTATA TTTTGGAAAA AGTAAAAGAA CACCAAGAAT CAATGGACAT GAACAACCCT
 781 CAGGACTTTA TTGATTGCTT CCTGATGAAA ATGGAGAAGG AAAAGCACAA CCAACCATCT
 841 GAATTTACTA TTGAAAGCTT GGAAACACT GCAGTTGACT TGTTTGGAGC TGGGACAGAG
 901 ACGACAAGCA CAACCCTGAG ATATGCTCTC CTTCTCCTGC TGAAGCACCC AGAGGTCACA
 961 GCTAAAGTCC AGGAAGAGAT TGAACGTGTG ATTGGCAGAA ACCGGAGCCC CTGCATGCAA
1021 GACAGGAGCC ACATGCCCTA CACAGATGCT GTGGTGCACG AGGTCCAGAG ATACATTGAC
1081 CTTCTCCCCA CCAGCCTGCC CCATGCAGTG ACCTGTGACA TTAAATTCAG AAACTATCTC
1141 ATTCCCAAGG GCACAACCAT ATTAATTTCC CTGACTTCTG TGCTACATGA CAACAAAGAA
1201 TTTCCCAACC CAGAGATGTT TGACCCTCAT CACTTTCTGG ATGAAGGTGG CAATTTTAAG
1261 AAAAGTAAAT ACTTCATGCC TTTCTCAGCA GGAAAACGGA TTTGTGTGGG AGAAGCCCTG
1321 GCCGGCATGG AGCTGTTTTT ATTCCTGACC TCCATTTTAC AGAACTTTAA CCTGAAATCT
1381 CTGGTTGACC CAAAGAACCT TGACACCACT CCAGTTGTCA ATGGATTTGC CTCTGTGCCG
1441 CCCTTCTACC AGCTGTGCTT CATTCCTGTC TGAAGAAGAG CAGATGGCCT GGCTGCTGCT
1501 GTGCAGTCCC TGCAGCTCTC TTTCCTCTGG GGCATTATCC ATCTTTGCAC TATCTGTAAT
1561 GCCTTTTCTC ACCTGTCATC TCACATTTTC CCTTCCCTGA AGATCTAGTG AACATTCGAC·
1621 CTCCATTACG GAGAGTTTCC TATGTTTCAC TGTGCAAATA TATCTGCTAT TCTCCATACT
1681 CTGTAACAGT TGCATTGACT GTCACATAAT GCTCATACTT ATCTAATGTA GAGTATTAAT
1741 ATGTTATTAT TAAATAGAGA AATATGATTT GTGTATTATA ATTCAAAGGC ATTTCTTTTC
1801 TGCATGATCT AAATAAAAAG CATTATTATT TGCTG
```

Figure 4A

```
   1 MDSLVVLVLC LSCLLLLSLW RQSSGRGKLP PGPTPLPVIG NILQIGIKDI SKSLTNLSKV
  61 YGPVFTLYFG LKPIVVLHGY EAVKEALIDL GEEFSGRGIF PLAERANRGF GIVFSNGKKW
 121 KEIRRFSLMT LRNFGMGKRS IEDRVQEEAR CLVEELRKTK ASPCDPTFIL GCAPCNVICS
 181 IIFHKRFDYK DQQFLNLMEK LNENIKILSS PWIQICNNFS PIIDYFPGTH NKLLKNVAFM
 241 KSYILEKVKE HQESMDMNNP QDFIDCFLMK MEKEKHNQPS EFTIESLENT AVDLFGAGTE
 301 TTSTTLRYAL LLLLKHPEVT AKVQEEIERV IGRNRSPCMQ DRSHMPYTDA VVHEVQRYID
 361 LLPTSLPHAV TCDIKFRNYL IPKGTTILIS LTSVLHDNKE FPNPEMFDPH HFLDEGGNFK
 421 KSKYFMPFSA GKRICVGEAL AGMELFLFLT SILQNFNLKS LVDPKNLDTT PVVNGFASVP
 481 PFYQLCFIPV *RRADGLAAA VQSLQLSFLW GIIHLCTICN AFSHLSSHIF PSLKI**TFD
 541 LHYGEFPMFH CANISAILHT L*QLH*LSHN AHTYLM*SIN MLLLNREI*F VYYNSKAFLF
 601 CMI*IKSIII C
```

Figure 4B

```
   1 ATGGGGCTAG AAGCACTGGT GCCCCTGGCC GTGATAGTGG CCATCTTCCT GCTCCTGGTG
  61 GACCTGATGC ACCGGCGCCA ACGCTGGGCT GCACGCTACC CACCAGGCCC CCTGCCACTG
 121 CCCGGGCTGG GCAACCTGCT GCATGTGGAC TTCCAGAACA CACCATACTG CTTCGACCAG
 181 TTGCGGCGCC GCTTCGGGGA CGTGTTCAGC CTGCAGCTGG CCTGGACGCC GGTGGTCGTG
 241 CTCAATGGGC TGGCGGCCGT GCGCGAGGCG CTGGTGACCC ACGGCGAGGA CACCGCCGAC
 301 CGCCCGCCTG TGCCCATCAC CCAGATCCTG GGTTTCGGGC CGCGTTCCCA AGGGGTGTTC
 361 CTGGCGCGCT ATGGGCCCGC GTGGCGCGAG CAGAGGCGCT TCTCCGTGTC CACCTTGCGC
 421 AACTTGGGCC TGGGCAAGAA GTCGCTGGAG CAGTGGGTGA CCGAGGAGGC CGCCTGCCTT
 481 TGTGCCGCCT TCGCCAACCA CTCCGGACGC CCCTTTCGCC CCAACGGTCT CTTGGACAAA
 541 GCCGTGAGCA ACGTGATCGC CTCCCTCACC TGCGGGCGCC GCTTCGAGTA CGACGACCCT
 601 CGCTTCCTCA GGCTGCTGGA CCTAGCTCAG GAGGGACTGA AGGAGGAGTC GGGCTTTCTG
 661 CGCGAGGTGC TGAATGCTGT CCCCGTCCTC CTGCATATCC CAGCGCTGGC TGGCAAGGTC
 721 CTACGCTTCC AAAAGGCTTT CCTGACCCAG CTGGATGAGC TGCTAACTGA GCACAGGATG
 781 ACCTGGGACC CAGCCCAGCC CCCCCGAGAC CTGACTGAGG CCTTCCTGGC AGAGATGGAG
 841 AAGGCCAAGG GGAACCCTGA GAGCAGCTTC AATGATGAGA ACCTGCGCAT AGTGGTGGCT
 901 GACCTGTTCT CTGCCGGGAT GGTGACCACC TCGACCACGC TGGCCTGGGG CCTCCTGCTC
 961 ATGATCCTAC ATCCGGATGT GCAGCGCCGT GTCCAACAGG AGATCGACGA CGTGATAGGG
1021 CAGGTGCGGC GACCAGAGAT GGGTGACCAG GCTCACATGC CCTACACCAC TGCCGTGATT
1081 CATGAGGTGC AGCGCTTTGG GGACATCGTC CCCCTGGGTA TGACCCATAT GACATCCCGT
1141 GACATCGAAG TACAGGGCTT CCGCATCCCT AAGGGAACGA CACTCATCAC CAACCTGTCA
1201 TCGGTGCTGA AGGATGAGGC CGTCTGGGAG AAGCCCTTCC GCTTCCACCC CGAACACTTC
1261 CTGGATGCCC AGGGCCACTT TGTGAAGCCG GAGGCCTTCC TGCCTTTCTC AGCAGGCCGC
1321 CGTGCATGCC TCGGGGAGCC CCTGGCCCGC ATGGAGCTCT TCCTCTTCTT CACCTCCCTG
1381 CTGCAGCACT TCAGCTTCTC GGTGCCCACT GGACAGCCCC GGCCCAGCCA CCATGGTGTC
1441 TTTGCTTTCC TGGTGAGCCC ATCCCCCTAT GAGCTTTGTG CTGTGCCCCG CTAG
```

Figure 5A

```
   1 MGLEALVPLA VIVAIFLLLV DLMHRRQRWA ARYPPGPLPL PGLGNLLHVD FQNTPYCFDQ
  61 LRRRFGDVFS LQLAWTPVVV LNGLAAVREA LVTHGEDTAD RPPVPITQIL GFGPRSQGVF
 121 LARYGPAWRE QRRFSVSTLR NLGLGKKSLE QWVTEEAACL CAAFANHSGR PFRPNGLLDK
 181 AVSNVIASLT CGRRFEYDDP RFLRLLDLAQ EGLKEESGFL REVLNAVPVL LHIPALAGKV
 241 LRFQKAFLTQ LDELLTEHRM TWDPAQPPRD LTEAFLAEME KAKGNPESSF NDENLRIVVA
 301 DLFSAGMVTT STTLAWGLLL MILHPDVQRR VQQEIDDVIG QVRRPEMGDQ AHMPYTTAVI
 361 HEVQRFGDIV PLGMTHMTSR DIEVQGFRIP KGTTLITNLS SVLKDEAVWE KPFRFHPEHF
 421 LDAQGHFVKP EAFLPFSAGR RACLGEPLAR MELFLFFTSL LQHFSFSVPT GQPRPSHHGV
 481 FAFLVSPSPY ELCAVPR*
```
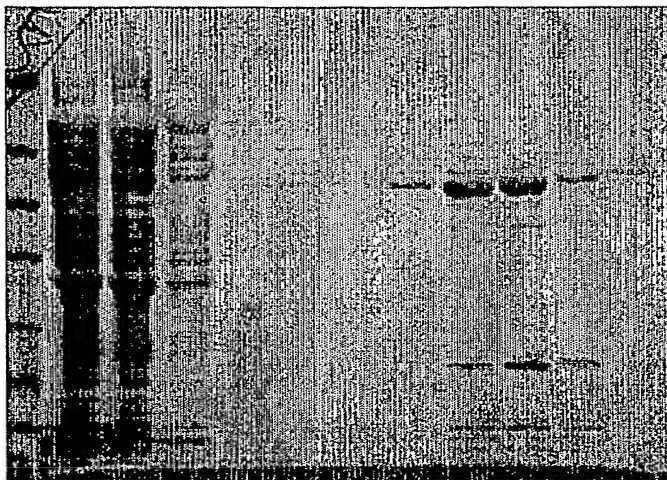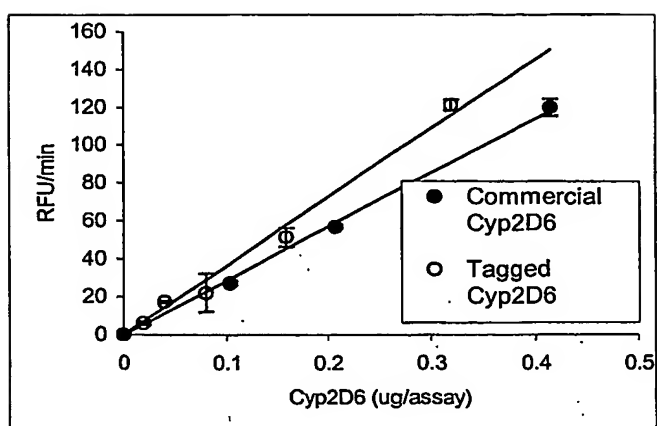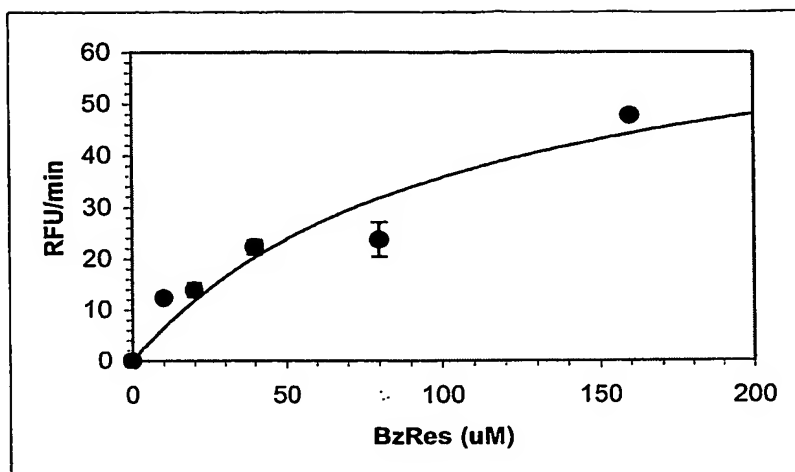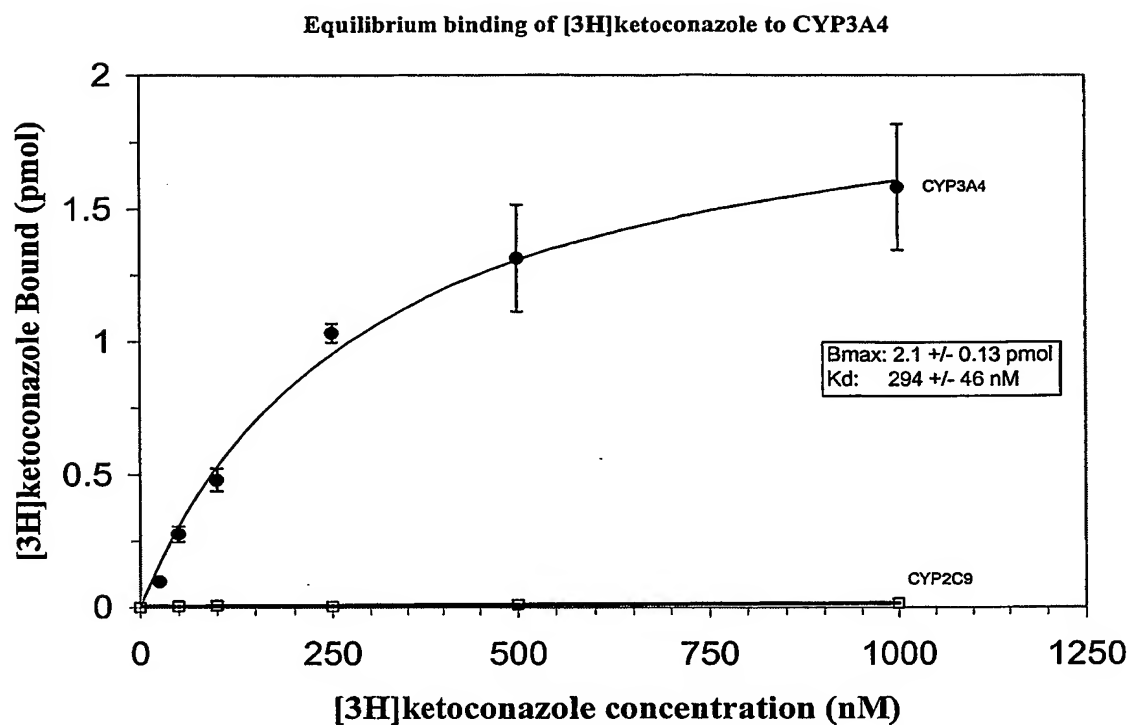
Figure 5B

Lane1   2   3   4   5   6   7   8   1   2   3   4   5   6   7   8
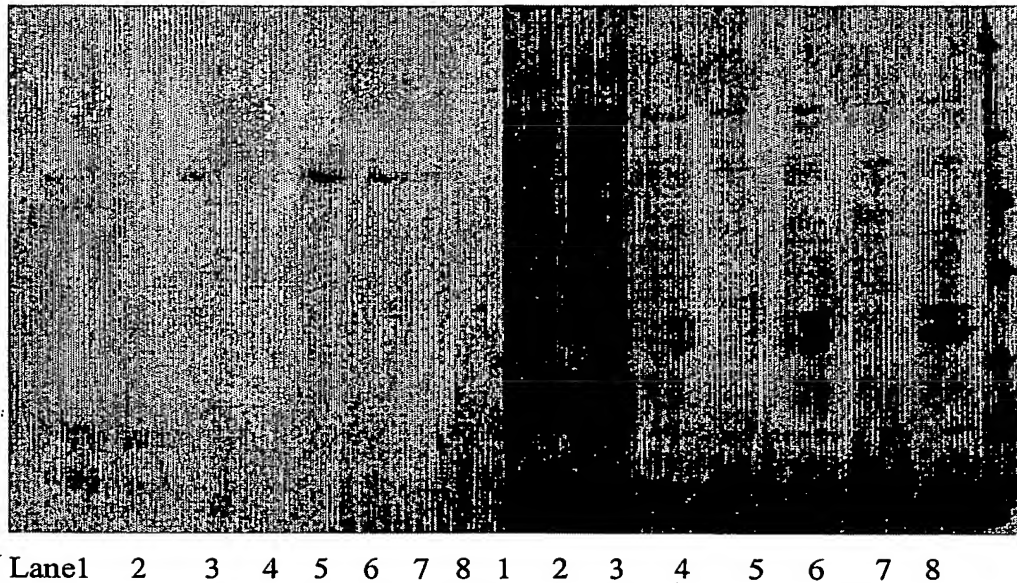
Figure 6

Figure 7

Figure 8

Figure 9

Equilibrium binding of [3H]ketoconazole to CYP3A4

Figure 10

Conversion of DBF to Fluorescein by Tagged Immobilised P450 3A4
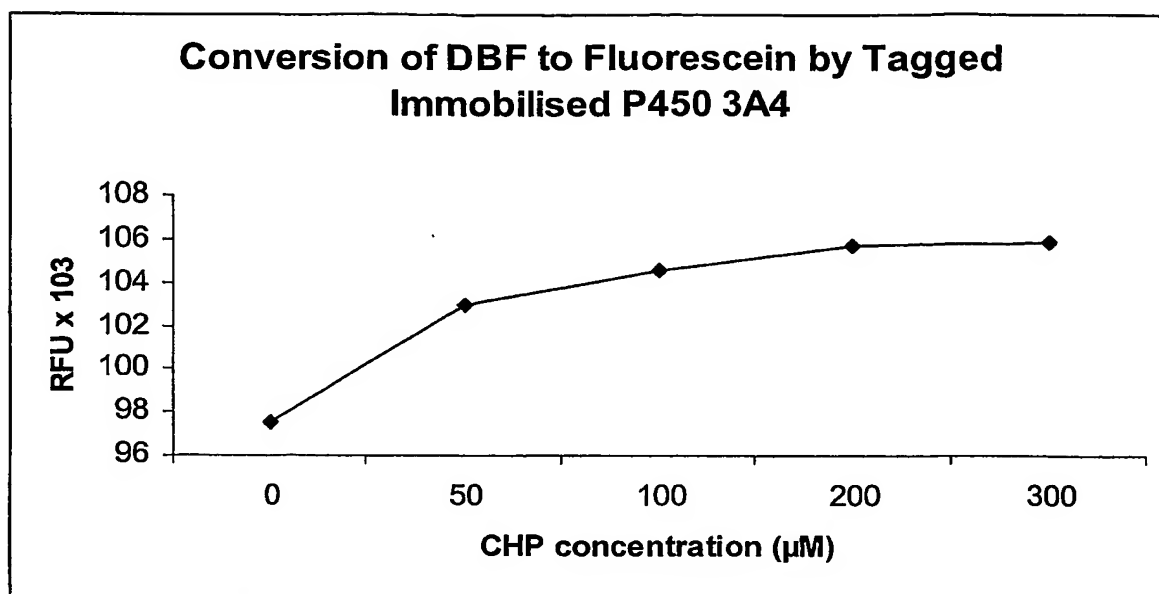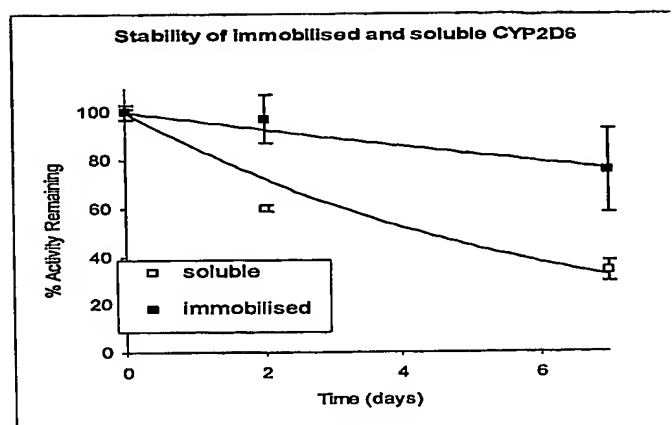
Figure 11

Figure 12

Figure 13

Figure 14